On the Achievable Rate Regions for Interference Channels with Degraded Message Sets[†]

Jinhua Jiang and Yan Xin
Department of Electrical and Computer Engineering
National University of Singapore, Singapore 117576
Email: { jinhua.jiang, elexy }@nus.edu.sg

Abstract

The interference channel with degraded message sets (IC-DMS) refers to a communication model in which two senders attempt to communicate with their respective receivers simultaneously through a common medium, and one of the senders has *complete* and *a priori* (non-causal) knowledge about the message being transmitted by the other. A coding scheme that collectively has advantages of cooperative coding, collaborative coding, and dirty paper coding, is developed for such a channel. With resorting to this coding scheme, achievable rate regions of the IC-DMS in both discrete memoryless and Gaussian cases are derived, which, in general, include several previously known rate regions. Numerical examples for the Gaussian case demonstrate that in the *high-interference-gain* regime, the derived achievable rate regions offer considerable improvements over these existing results.

Index Terms

Cognitive radio, cooperative communication, degrade message sets, dirty paper coding, Gel'fand-Pinsker coding, interference channels, superposition coding.

[†] The work is supported by the National University of Singapore (NUS) under start-up grants R-263-000-314-101 and R-263-000-314-112 and by a NUS Research Scholarship. The correspondence author of the paper is Dr. Yan Xin (tel. no. +65 6516-5513 and fax no. +65 6779-1103).

I. INTRODUCTION

The interference channel with degraded message sets (IC-DMS) refers a communication model in which two senders attempt to communicate with their respective receivers simultaneously through a common medium, and one of the senders has *complete* and *a priori* (non-causal) knowledge about the message being transmitted by the other. Such a model generically characterizes some realistic communication scenarios taking place in cognitive radio channels [1], [2] or in wireless sensor networks over a correlated field [3], [4], which we illustrate in Figs. 1(a) and 1(b).

From an information-theoretic perspective, the IC-DMS have been investigated in [1]–[4]. Specifically, several achievable rate results have been obtained in [1]–[4], and the capacity regions for two special cases have been characterized in [2]–[4]. The main achievable rate region in [1] was obtained by incorporating the Gel'fand-Pinsker coding [5] into the well-known coding scheme applied to the interference channel (IC) [6], [7]. In this coding scheme, each of the two senders splits its message into two sub-messages, and allows its non-pairing receiver to decode one of the sub-messages. Knowing the two sub-messages and the corresponding codewords which sender 1 wishes to transmit, sender 2 applies the Gel'fand-Pinsker coding to encode its own sub-messages by treating the codewords of sender 1 as known interferences. It has been also shown in [1, Corollary 2] that, an improved achievable rate region can be attained by time-sharing between the early derived rate region and a so called fully-cooperative rate point achieved by letting sender 2 use all its power to transmit sender 1' messages. A different coding scheme was adopted in [2] and [3], in which neither of the senders splits its message into two sub-messages, and receiver 2 does not decode any transmitted information from sender 1. Since sender 2 knows what sender 1 wishes to transmit, sender 2 is allowed to: 1) apply the Gel'fand-Pinsker coding to encode its own message; and 2) partially cooperate with sender 1 using superposition coding. It has been proven in [2], [3] that, this is the capacity-achieving scheme for the Gaussian IC-DMS in the *low-interference-gain* regime, in which the normalized link gain between sender 2 and receiver 1 is less than or equal to 1.

However, in practice, due to the mobility of the users or random distributions of the sensors,

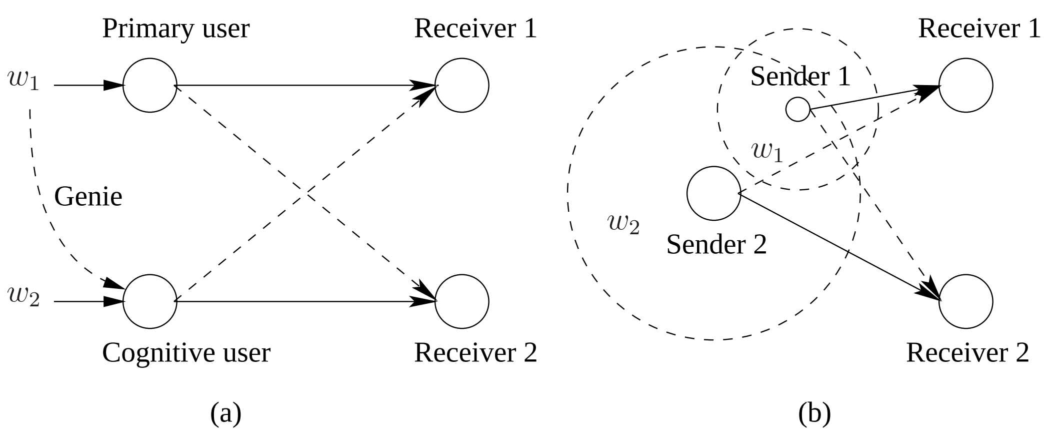

Fig. 1. (a) A genie-aided cognitive radio channel [1], in which the Genie informs the cognitive user of what the primary user will transmit; (b) A four-node wireless sensor network [3], in which sender 2 senses a larger area such that it knows what information sender 1 obtains.

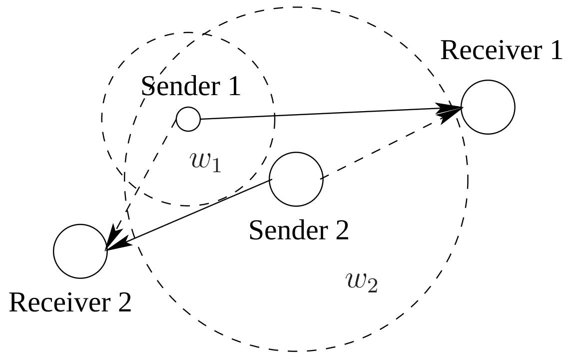

Fig. 2. An interference channel with degraded message sets in which sender 2 is close to receiver 1.

sender 2 may be geographically located near to receiver 1, as illustrated in Fig. 2. It is likely, in such a situation, that the Gaussian IC-DMS is in the *high-interference-gain* regime, in which the normalized link gain between sender 2 and receiver 1 is greater than 1. In fact, the findings in this paper reveal that the achievable rate region, which was proven to be the capacity region in the low-interference-gain regime in [2] and [3], is *strictly* non-optimal for the Gaussian IC-DMS in the high-interference-gain regime.

In this paper, we develop a new coding scheme for the IC-DMS to improve existing achievable rate regions. Our coding scheme differs from one proposed in [2], [3] in the way that, sender 2 splits its message into two sub-messages, and encodes both sub-messages using Gel'fand-Pinsker coding. Moreover, receiver 1 is required to jointly decode the message from sender 1 and one sub-message from sender 2. With this coding scheme, we derive our main achievable rate region for the discrete memoryless case. For comparison purpose, we compromise either the coding flexibility (fixing an auxiliary random variable as a constant), or the advantage of simultaneous decoding [7], to obtain two subregions of the main achievable rate region. The obtained subregions are shown to either include or be the same as the existing ones. We further extend the obtained regions from the discrete memoryless case to the Gaussian case, and show by numerical examples that our Gaussian achievable rate results strictly improve the existing ones in the high-interference-gain regime.

The rest of the paper is organized as follows. In Section II, we introduce the channel model of the IC-DMS, and the related terminologies. In Section III, we present the main achievable result for the discrete memoryless case with a detailed proof. In Section IV, we derive two subregions of the main achievable rate region, and we show that the derived subregions include several existing results as special cases. Lastly, in Section V, we extend our rate regions from the discrete memoryless case to the Gaussian case, and compare them with the existing results.

Notations: Random variables and their realizations are denoted by upper case letters and lower case letters respectively, e.g., X and x. Bold lower (upper) case letters are used to denote vectors (matrices), e.g., $\mathbf{x}$ and $\mathbf{\Sigma}$. Calligraphic fonts are used to denote sets, e.g., $\mathcal{X}$ and $\mathcal{R}$.

II. The Channel Model

Consider the IC-DMS (also termed as the genie-aided cognitive radio channel in [1]) depicted in Fig. 3, in which sender 1 wishes to transmit a message (or message index), $w_1 \in \mathcal{M}_1 := \{1, ..., M_1\}$, to receiver 1 and sender 2 wishes to transmit its message,

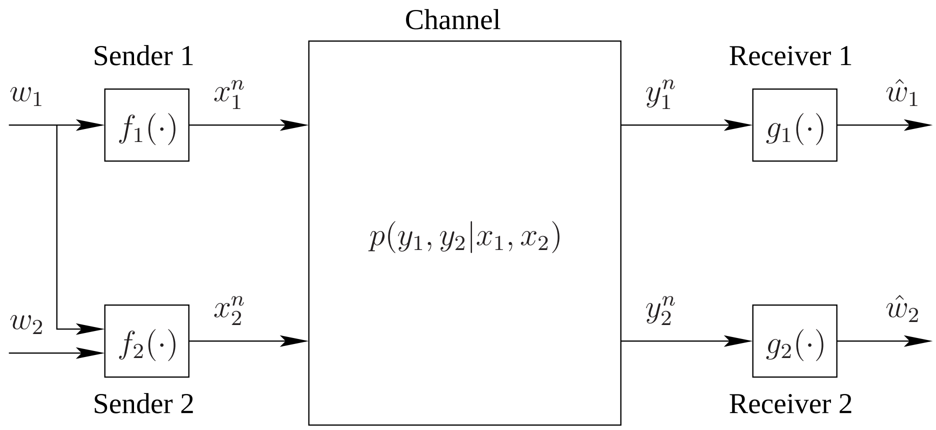

Fig. 3. An interference channel with degraded message sets.

$w_2 \in \mathcal{M}_2 := \{1, ..., M_2\}$, to receiver 2. Typically, this *discrete memoryless* IC-DMS is described by a tuple $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2, p(y_1, y_2|x_1, x_2))$, where $\mathcal{X}_1$ and $\mathcal{X}_2$ are the channel input alphabets, $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are the channel output alphabets, and $p(y_1, y_2|x_1, x_2)$ denotes the conditional probability of $(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ given $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. The channel is discrete memoryless in the sense that

$$p(y_{1,t}, y_{2,t}|x_{1,t}, x_{2,t}, x_{1,t-1}, x_{2,t-1}, ...) = p(y_{1,t}, y_{2,t}|x_{1,t}, x_{2,t}), \tag{1}$$

for every discrete time instant t in a synchronous transmission. In terms of the channel input-output relationship, the IC-DMS is the same as the IC. However, in the IC-DMS, sender 2 is able to noncausally obtain the knowledge of the message w_1, which will be transmitted from sender 1. This is the key difference between the IC-DMS and IC in terms of the information flow. We next present the following standard definitions with regard to the existence of codes and achievable rates for the discrete memoryless IC-DMS channel.

Definition 1: An (M_1, M_2, n, P_e) code exists for the discrete memoryless IC-DMS, if and only if there exist two encoding functions

$$f_1 : \mathcal{M}_1 \to \mathcal{X}_1^n, \quad f_2 : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{X}_2^n,$$

and two decoding functions

$$g_1 : \mathcal{Y}_1^n \to \mathcal{M}_1, \quad g_2 : \mathcal{Y}_2^n \to \mathcal{M}_2,$$

such that $\max\{P_{e,1}^{(n)}, P_{e,2}^{(n)}\} \leq P_e$, where $P_{e,1}^{(n)}$ and $P_{e,2}^{(n)}$ denote the respective average probabilities of error at decoders 1 and 2, and are computed as

$$P_{e,1}^{(n)} = \frac{1}{M_1 M_2} \sum_{w_1 w_2} p(\hat{w}_1 \neq w_1 | (w_1, w_2) \text{ were sent}),$$

$$P_{e,2}^{(n)} = \frac{1}{M_1 M_2} \sum_{w_1 w_2} p(\hat{w}_2 \neq w_2 | (w_1, w_2) \text{ were sent}).$$

Definition 2: A non-negative rate pair (R_1, R_2) is achievable for the IC-DMS, if for any given $0 < P_e < 1$ and any sufficiently large n, there exists a $(2^{nR_1}, 2^{nR_2}, n, P_e)$ code for the channel. The capacity region of the IC-DMS is the set of all the achievable rate pairs for the channel, and an achievable rate region is a subset of the capacity region.

It should be noted that from an information-theoretic standpoint, the IC can not be simply treated as a special case of the IC-DMS in the sense that the capacity region of the IC-DMS, if any, does not imply a capacity region of the IC.

III. An Achievable Rate Region for the Discrete Memoryless IC-DMS

In this section, we present the main achievable rate region for the discrete memoryless IC-DMS, which is the primary result in this paper.

Consider auxiliary random variables W, U, $\tilde{U}$, V, $\tilde{V}$ and a time-sharing random variable Q, defined on arbitrary finite sets $\mathcal{W}$, $\mathcal{U}$, $\tilde{\mathcal{U}}$, $\mathcal{V}$, $\tilde{\mathcal{V}}$ and $\mathcal{Q}$ respectively. Let $\mathcal{P}$ denote the set of all joint probability distributions $p(\cdot)$ that factor in the form of

$$\begin{aligned} p(q, w, x_1, u, \tilde{u}, v, \tilde{v}, x_2, y_1, y_2) =& p(q)p(w, x_1|q)p(u, \tilde{u}|w, q)p(v, \tilde{v}|w, q) \\ & \cdot p(x_2|\tilde{u}, \tilde{v}, w, q)p(y_1, y_2|x_1, x_2), \end{aligned} \tag{2}$$

where w, u, $\tilde{u}$, v, $\tilde{v}$, and q are realizations of random variables W, U, $\tilde{U}$, V, $\tilde{V}$ and Q.

Let $\mathcal{R}(p)$ denote the set of all non-negative rate pairs (R_1, R_2) such that the following inequalities hold simultaneously

$$R_1 \le I(W; Y_1 U|Q), \tag{3}$$
$$R_2 \le I(UV; Y_2|Q) - I(U; W|Q) - I(V; W|Q), \tag{4}$$
$$R_1 + R_2 \le I(UW; Y_1|Q) + I(V; Y_2 U|Q) - I(U; W|Q) - I(V; W|Q); \tag{5}$$
$$0 \le I(UW; Y_1|Q) - I(U; W|Q), \tag{6}$$
$$0 \le I(U; Y_2 V|Q) - I(U; W|Q), \tag{7}$$
$$0 \le I(V; Y_2 U|Q) - I(V; W|Q), \tag{8}$$
$$0 \le I(UV; Y_2|Q) - I(U; W|Q) - I(V; W|Q), \tag{9}$$

for a given joint distribution $p(\cdot) \in \mathcal{P}$.

Let $\mathcal{C}$ denote the capacity region of the discrete memoryless IC-DMS, and let

$$\mathcal{R} = \bigcup_{p(\cdot)\in\mathcal{P}} \mathcal{R}(p).$$

Theorem 1: The region $\mathcal{R}$ is an achievable rate region for the discrete memoryless IC-DMS, i.e., $\mathcal{R} \subseteq \mathcal{C}$.

Proof: Before presenting the proof of Theorem 1, we state the following lemma as it will be frequently used in the proof.

Lemma 1 ([8, Theorem 14.2.3]): Let $A_\epsilon^{(n)}$ denote the typical set for the probability mass distribution $p(s_1, s_2, s_3)$, and let $P(\mathbf{S}'_1 = \mathbf{s}_1, \mathbf{S}'_2 = \mathbf{s}_2, \mathbf{S}'_3 = \mathbf{s}_3) = \prod_{i=1}^n p(s_{1i}|s_{3i})p(s_{2i}|s_{3i})p(s_{3i})$, then $P\{(\mathbf{S}'_1, \mathbf{S}'_2, \mathbf{S}'_3) \in A_\epsilon^{(n)}\} \doteq 2^{-n(I(S'_1;S'_2|S'_3)\pm 6\epsilon)}$.

To prove this theorem we apply the notion of the asymptotic equipartition property (APE) [8]. Our coding scheme is mainly based on the arguments of superposition coding [9] and

Gel'fand-Pinsker coding [5]. Specifically, sender 1 independently encodes its message w_1 as a whole; while sender 2 needs split its message into two parts, i.e., $w_2 = (w_{21}, w_{22})$, and encode them separately. Both w_{21} and w_{22} are encoded using the Gel'fand-Pinsker approach, but they are processed differently at the receivers. The message w_{22} will be decoded by receiver 2 only, while w_{21} will be decoded by both receivers. Moreover, knowing the message and codeword which sender 1 is going to transmit, sender 2 not only can apply Gel'fand-Pinsker coding to deal with the known interference, but also can cooperate with sender 1 to transmit w_1 using superposition coding. Let R_{21} and R_{22} denote the rates of w_{21} and w_{22} respectively, i.e., $w_{21} \in \{1, \ldots, 2^{nR_{21}}\}$ and $w_{22} \in \{1, \ldots, 2^{nR_{22}}\}$. If receiver 1 can decode w_1 and receiver 2 can decode both w_{21} and w_{22} with vanishing probabilities of error, then $(R_1, R_{21} + R_{22})$ is an achievable rate pair for the IC-DMS.

To prove that the entire region $\mathcal{R}$ is achievable for the channel, it is sufficient to prove that $\mathcal{R}(p)$ is achievable for a fixed joint probability distribution $p(\cdot) \in \mathcal{P}$.

A. Random Codebook Generation

Consider a fixed joint distribution $p(\cdot) \in \mathcal{P}$, and a random time-sharing codeword $\mathbf{q}$ of length n, which is given to both senders and receivers. The codeword $\mathbf{q}$ is assumed to be generated according to $\prod_{i=1}^n p(q_i)$.

Generate 2^{nR_1} independent codewords $\mathbf{w}(j)$, $j \in \{1, \ldots, 2^{nR_1}\}$, according to $\prod_{i=1}^n p(w_i|q_i)$; and for each $\mathbf{w}(j)$ generate one $\mathbf{x}_1(j)$, according to $\prod_{i=1}^n p(x_{1i}|w_i q_i)$. Similarly, generate $2^{n\tilde{R}_{21}}$ independent codewords $\mathbf{u}(l_1)$, $l_1 \in \{1, \ldots, 2^{n\tilde{R}_{21}}\}$, according to $\prod_{i=1}^n p(u_i|q_i)$, and generate $2^{n\tilde{R}_{22}}$ independent codewords $\mathbf{v}(l_2)$, $l_2 \in \{1, \ldots, 2^{n\tilde{R}_{22}}\}$, according to $\prod_{i=1}^n p(v_i|q_i)$.

For each codeword pair $(\mathbf{u}(l_1), \mathbf{w}(j))$, generate one codeword $\tilde{\mathbf{u}}(l_1, j)$ according to $\prod_{i=1}^n p(\tilde{u}_i|u_i(l_1) w_i(j) q_i)$, and similarly for each codeword pair $(\mathbf{v}(l_2), \mathbf{w}(j))$, generate one codeword $\tilde{\mathbf{v}}(l_2, j)$ according to $\prod_{i=1}^n p(\tilde{v}_i|v_i(l_2) w_i(j) q_i)$. Lastly, for each codeword triple $(\mathbf{u}(l_1), \mathbf{v}(l_2), \mathbf{w}(j))$, generate one codeword $\mathbf{x}_2(l_1, l_2, j)$ according to $\prod_{i=1}^n p(x_{2i}|\tilde{u}_i(l_1) \tilde{v}_i(l_2) w_i(j) q_i)$.

Now uniformly distribute $2^{n\tilde{R}_{21}}$ codewords $\mathbf{u}(l_1)$ into $2^{nR_{21}}$ bins indexed by $k_1 \in \{1, \ldots, 2^{nR_{21}}\}$ such that each bin contains $2^{n(\tilde{R}_{21} - R_{21})}$ codewords; uniformly distribute $2^{n\tilde{R}_{22}}$ codewords $\mathbf{v}(l_2)$ into $2^{nR_{22}}$ bins indexed by $k_2 \in \{1, \ldots, 2^{nR_{22}}\}$ such that each bin contains $2^{n(\tilde{R}_{22} - R_{22})}$ codewords.

The entire codebook is revealed to both senders and receivers.

B. Encoding and Transmission

We assume that the senders want to transmit a message vector $(w_1, w_{21}, w_{22}) = (j, k_1, k_2)$. Sender 1 simply encodes the message as codeword $\mathbf{x}_1(j)$ and sends the codeword with n channel uses. Sender 2 will first need to look for a codeword $\mathbf{u}(\hat{l}_1)$ in bin k_1 such that $(\mathbf{u}(\hat{l}_1), \mathbf{w}(j), \mathbf{q}) \in A_\epsilon^{(n)}$, and a codeword $\mathbf{v}(\hat{l}_2)$ in bin k_2 such that $(\mathbf{v}(\hat{l}_2), \mathbf{w}(j), \mathbf{q}) \in A_\epsilon^{(n)}$. If sender 2 fails to do so, it will randomly pick a codeword $\mathbf{u}(\hat{l}_1)$ from bin k_1 or a codeword $\mathbf{v}(\hat{l}_2)$ from bin k_2. Sender 2 then transmits codeword $\mathbf{x}_2(\hat{l}_1, \hat{l}_2, j)$ through n channel uses. We further assume that the transmissions are perfectly synchronized.

C. Decoding

Receiver 1 first looks for all the index pairs $(\hat{j}, \hat{\hat{l}}_1)$ such that $(\mathbf{w}(\hat{j}), \mathbf{u}(\hat{\hat{l}}_1), \mathbf{y}_1, \mathbf{q}) \in A_\epsilon^{(n)}$. If $\hat{j}$ in all the index pairs found are the same, receiver 1 determines $w_1 = \hat{j}$, otherwise declares an error.

Receiver 2 will first look for all index pairs $(\bar{\hat{l}}_1, \hat{\hat{l}}_2)$ such that $(\mathbf{u}(\bar{\hat{l}}_1), \mathbf{v}(\hat{\hat{l}}_2), \mathbf{y}_2, \mathbf{q}) \in A_\epsilon^{(n)}$. If $\bar{\hat{l}}_1$ in all the index pairs found are indices of codewords $\mathbf{u}(\bar{\hat{l}}_1)$ from the same bin with index $\hat{k}_1$, and $\hat{\hat{l}}_2$ in all the index pairs found are indices of codewords $\mathbf{v}(\hat{\hat{l}}_2)$ from the same bin with index $\hat{k}_2$, then receiver 2 will decode that $(w_{21}, w_{22}) = (\hat{k}_1, \hat{k}_2)$ were transmitted; otherwise, an error is declared.

D. Evaluation of Probability of Error

We now derive upper bounds for the probabilities of the respective error events, which may happen during the encoding and decoding process. Due to the symmetry of the codebook generation and encoding processing, the probability of error is not codeword dependent. Without loss of generality, we assume that $(w_1, w_{21}, w_{22}) = (1, 1, 1)$ were encoded and transmitted. We next define the following three types of events:

$$\begin{aligned}
E_{a,b} &= (\mathbf{u}(a), \mathbf{w}(b), \mathbf{q}) \in A_\epsilon^{(n)}, \\
\dot{E}_{a,b} &= (\mathbf{w}(a), \mathbf{u}(b), \mathbf{y}_1, \mathbf{q}) \in A_\epsilon^{(n)}, \\
\ddot{E}_{a,b} &= (\mathbf{u}(a), \mathbf{v}(b), \mathbf{y}_2, \mathbf{q}) \in A_\epsilon^{(n)}.
\end{aligned}$$

Let $P_e(\text{enc2})$, $P_e(\text{dec1})$, and $P_e(\text{dec2})$ denote the probabilities of error at the encoder of sender 2, the decoder of receiver 1, and the decoder of receiver 2, respectively.

[**Evaluation of** $P_e(\text{enc2})$.] An error is made if 1) the encoder at sender 2 can not find $\mathbf{u}(\hat{l}_1)$ in bin 1 such that $(\mathbf{u}(\hat{l}_1), \mathbf{w}(1), \mathbf{q}) \in A_\epsilon^{(n)}$, and/or 2) it can not find $\mathbf{v}(\hat{l}_2)$ in bin 1 such that $(\mathbf{v}(\hat{l}_2), \mathbf{w}(1), \mathbf{q}) \in A_\epsilon^{(n)}$. Then the probability of error at the encoder of sender 2 is bounded as

$$\begin{aligned}
P_e(\text{enc2}) &\leq Pr\left(\bigcap_{\mathbf{u}(\hat{l}_1)\in\text{bin } 1} (\mathbf{u}(\hat{l}_1), \mathbf{w}(1), \mathbf{q}) \notin A_\epsilon^{(n)}\right) + Pr\left(\bigcap_{\mathbf{v}(\hat{l}_2)\in\text{bin } 1} (\mathbf{v}(\hat{l}_2), \mathbf{w}(1), \mathbf{q}) \notin A_\epsilon^{(n)}\right) \\
&= \prod_{\mathbf{u}(\hat{l}_1)\in\text{bin } 1} Pr(E^c_{\hat{l}_1,1}) + \prod_{\mathbf{v}(\hat{l}_2)\in\text{bin } 1} Pr(E^c_{\hat{l}_2,1}) \\
&\leq (1 - Pr(E_{\hat{l}_1,1}))^{2^{n(\tilde{R}_{21}-R_{21})}} + (1 - Pr(E_{\hat{l}_2,1}))^{2^{n(\tilde{R}_{22}-R_{22})}} \\
&\overset{(a)}{\leq} (1 - 2^{-n(I(U;W|Q)+6\epsilon)})^{2^{n(\tilde{R}_{21}-R_{21})}} + (1 - 2^{-n(I(V;W|Q)+6\epsilon)})^{2^{n(\tilde{R}_{22}-R_{22})}},
\end{aligned}$$

where (a) follows from the fact that we can obtain $Pr(E_{\hat{l}_1,1}) \geq 2^{-n(I(U;W|Q)+6\epsilon)}$ and $Pr(E_{\hat{l}_2,1}) \geq 2^{-n(I(V;W|Q)+6\epsilon)}$ by setting $\mathbf{S}'_1 = \mathbf{U}$, $\mathbf{S}'_2 = W$, and $\mathbf{S}'_3 = \mathbf{Q}$, and $\mathbf{S}'_1 = \mathbf{V}$, $\mathbf{S}'_2 = W$, and $\mathbf{S}'_3 = \mathbf{Q}$ in Lemma 1, respectively. Following the same argument in the proof of Lemma 2.1.3 of [10], we conclude that $P_e(\text{enc2}) \to 0$ as $n \to +\infty$, if

$$\tilde{R}_{21} \geq R_{21} + I(U;W|Q), \tag{10}$$

$$\tilde{R}_{22} \geq R_{22} + I(V;W|Q), \tag{11}$$

are satisfied. We further choose

$$\tilde{R}_{21} = R_{21} + I(U;W|Q), \tag{12}$$
$$\tilde{R}_{22} = R_{22} + I(V;W|Q). \tag{13}$$

Note that such a choice still ensures that $P_e(\text{enc2}) \to 0$ as $n \to +\infty$.

[**Evaluation of** $P_e(\text{dec1})$] An error is made if 1) $\dot{E}^c_{1,\hat{l}_1}$ happens, and/or 2) there exists some $\hat{j} \neq 1$ such that $\dot{E}_{\hat{j},\hat{\hat{l}}_1}$ happens. Note that $\hat{\hat{l}}_1$ is not required to be equal to $\hat{l}_1$, since it is unnecessary for receiver 1 to decode $\hat{l}_1$ correctly. The probability of error at receiver 1 can be upper bounded as

$$\begin{aligned}
P_e(\text{dec1}) &\leq Pr(\dot{E}^c_{1,\hat{l}_1} \bigcup \cup_{\hat{j}\neq 1} \dot{E}_{\hat{j},\hat{\hat{l}}_1}) \\
&\leq Pr(\dot{E}^c_{1,\hat{l}_1}) + \sum_{\hat{j}\neq 1} Pr(\dot{E}_{\hat{j},\hat{\hat{l}}_1}) \\
&= Pr(\dot{E}^c_{1,\hat{l}_1}) + \sum_{\hat{j}\neq 1} Pr(\dot{E}_{\hat{j},\hat{l}_1}) + \sum_{\hat{j}\neq 1, \hat{\hat{l}}_1 \neq \hat{l}_1} P(\dot{E}_{\hat{j},\hat{\hat{l}}_1}) \\
&\leq Pr(\dot{E}^c_{1,\hat{l}_1}) + 2^{nR_1} Pr(\dot{E}_{2,\hat{l}_1}) + 2^{n(R_1+\tilde{R}_{21})} Pr(\dot{E}_{2,\hat{\hat{l}}_1 \neq \hat{l}_1}).
\end{aligned} \tag{14}$$

Choosing $\mathbf{S}'_1 = \mathbf{W}$, $\mathbf{S}'_2 = (\mathbf{Y}_1, \mathbf{U})$, and $\mathbf{S}'_3 = \mathbf{Q}$ in Lemma 1, we have $Pr(\dot{E}_{2,\hat{l}_1}) \doteq 2^{-n(I(W;Y_1U|Q)\pm 6\epsilon)}$. Likewise, we have $Pr(\dot{E}_{2,\hat{\hat{l}}_1\neq\hat{l}_1}) \doteq 2^{-n(I(WU;Y_1|Q)\pm 6\epsilon)}$. In addition, it follows from AEP that $Pr(\dot{E}^c_{1,\hat{l}_1}) \to 0$ as $n \to +\infty$. Thus, we infer from (14) that $P_e(\text{dec1}) \to 0$ as $n \to +\infty$, if

$$R_1 \leq I(W;Y_1U|Q), \tag{15}$$
$$R_1 + \tilde{R}_{21} \leq I(WU;Y_1|Q), \tag{16}$$

are satisfied.

[**Evaluation of** $P_e(\text{dec2})$] An error is made if 1) $\ddot{E}^c_{\hat{l}_1,\hat{l}_2}$ happens, and/or 2) there exists some $(\bar{\hat{l}}_1, \hat{\hat{l}}_2)$ in which either $\bar{\hat{l}}_1$ or $\hat{\hat{l}}_2$ is not an index of any codeword from the respective bin 1. The probability of the second case is upper bounded by the probability of the event, $\ddot{E}_{\bar{\hat{l}}_1,\hat{\hat{l}}_2}$ for some $(\bar{\hat{l}}_1, \hat{\hat{l}}_2) \neq (\hat{l}_1, \hat{l}_2)$. Thus, the probability of error at receiver 2 is bounded as

$$\begin{aligned}
P_e(\text{dec2}) &\leq Pr(\ddot{E}^c_{\hat{l}_1,\hat{l}_2} \bigcup \cup_{(\bar{\hat{l}}_1,\hat{\hat{l}}_2)\neq(\hat{l}_1,\hat{l}_2)} \ddot{E}_{\bar{\hat{l}}_1,\hat{\hat{l}}_2}) \\
&\leq Pr(\ddot{E}^c_{\hat{l}_1,\hat{l}_2}) + \sum_{(\bar{\hat{l}}_1,\hat{\hat{l}}_2)\neq(\hat{l}_1,\hat{l}_2)} P(\ddot{E}_{\bar{\hat{l}}_1,\hat{\hat{l}}_2}) \\
&= Pr(\ddot{E}^c_{\hat{l}_1,\hat{l}_2}) + \sum_{\bar{\hat{l}}_1\neq\hat{l}_1} Pr(\ddot{E}_{\bar{\hat{l}}_1,\hat{l}_2}) + \sum_{\hat{\hat{l}}_2\neq\hat{l}_2} Pr(\ddot{E}_{\hat{l}_1,\hat{\hat{l}}_2}) + \sum_{(\bar{\hat{l}}_1\neq\hat{l}_1,\hat{\hat{l}}_2\neq\hat{l}_2)} Pr(\ddot{E}_{\bar{\hat{l}}_1,\hat{\hat{l}}_2}) \\
&\leq Pr(\ddot{E}^c_{\hat{l}_1,\hat{l}_2}) + 2^{n\tilde{R}_{21}} Pr(\ddot{E}_{\bar{\hat{l}}_1\neq\hat{l}_1,\hat{l}_2}) + 2^{n\tilde{R}_{22}} Pr(\ddot{E}_{\hat{l}_1,\hat{\hat{l}}_2\neq\hat{l}_2}) + 2^{n(\tilde{R}_{21}+\tilde{R}_{22})} Pr(\ddot{E}_{\bar{\hat{l}}_1\neq\hat{l}_1,\hat{\hat{l}}_2\neq\hat{l}_2}).
\end{aligned} \tag{17}$$

Applying Lemma 1 to evaluate $Pr(\ddot{E}_{\bar{l}_1\neq\hat{l}_1,\hat{l}_2})$, $Pr(\ddot{E}_{\hat{l}_1,\hat{\hat{l}}_2\neq\hat{l}_2})$ and $Pr(\ddot{E}_{\bar{l}_1\neq\hat{l}_1,\hat{\hat{l}}_2\neq\hat{l}_2})$ in (17), we conclude that $P_e(\text{dec2}) \to 0$ as $n \to +\infty$ if the following inequalities,

$$\tilde{R}_{21} \leq I(U;Y_2V|Q), \tag{18}$$

$$\tilde{R}_{22} \leq I(V;Y_2U|Q), \tag{19}$$

$$\tilde{R}_{21} + \tilde{R}_{22} \leq I(UV;Y_2|Q), \tag{20}$$

are satisfied.

According to (12), (13) and the fact that $R_2 = R_{21} + R_{22}$, we first substitute $\tilde{R}_{21}$ and $\tilde{R}_{22}$ with $R_{21} + I(U;W|Q)$ and $R_{22} + I(V;W|Q)$ in (15), (16) and (18)–(20), and subsequently substitute R_{21} with $R_2 - R_{22}$ in the resulting inequalities. After these two substitution steps, we have

$$R_1 \leq I(W;Y_1U|Q), \tag{21}$$

$$R_1 + R_2 - R_{22} \leq I(WU;Y_1|Q) - I(U;W|Q), \tag{22}$$

$$R_2 - R_{22} \leq I(U;Y_2V|Q) - I(U;W|Q), \tag{23}$$

$$R_{22} \leq I(V;Y_2U|Q) - I(V;W|Q), \tag{24}$$

$$R_2 \leq I(UV;Y_2|Q) - (I(U;W|Q) + I(V;W|Q)). \tag{25}$$

Furthermore, applying Fourier-Motzkin elimination [11] to remove R_{22} from (21)–(25), we have

$$R_1 \leq I(W;Y_1U|Q), \tag{26}$$

$$R_2 \leq I(UV;Y_2|Q) - (I(U;W|Q) + I(V;W|Q)), \tag{27}$$

$$R_2 \leq I(U;Y_2V|Q) - I(U;W|Q) + I(V;Y_2U|Q) - I(V;W|Q), \tag{28}$$

$$R_1 + R_2 \leq I(WU;Y_1|Q) - I(U;W|Q) + I(V;Y_2U|Q) - I(V;W|Q). \tag{29}$$

Since $I(U;Y_2V|Q) + I(V;Y_2U|Q) - I(UV;Y_2|Q) = I(U;V|Q) + I(U;V|Y_2Q) \geq 0$, (27) implies (28) and thus (28) is redundant. To ensure that R_1, R_{21} and R_{22} are non-negative, we enforce four additional constraints (6)–(9). Therefore, the rate region $\mathcal{R}(p)$ is achievable for a fixed joint probability distribution $p(\cdot) \in \mathcal{P}$, and Theorem 1 follows. ■

Remark 1: The proposed coding scheme exploits three coding methods to achieve any rate pair in the rate region, $\mathcal{R}$. The first method is *cooperation* that is realized by the superposition relationship between $\mathbf{w}$ and $\mathbf{x}_2$ through $p(x_2|\tilde{u}_2, \tilde{v}_2, w, q)$. The second is *collaboration*, by which we mean that sender 2 separates its own message into two parts, i.e., $w_2 = (w_{21}, w_{22})$, and encodes w_{21} at a possibly low rate such that receiver 1 can decode it. By doing so, the effective interference caused by the signals carrying the sender 2's information may be reduced. The third is *Gel'fand-Pinsker coding*, which we apply to encode both messages, w_{21} and w_{22}, from sender 2 by treating the codeword $\mathbf{w}$ as known interference. This perhaps allows receiver 2 to be able to decode the messages from sender 2 at the same rate as if the interference caused by sender 1 was not present [12].

IV. Relating with Existing Rate Regions

In this section, we will show that Theorem 1 includes the achievable rate regions in [2], [3]. To demonstrate it, we compromise the advantages of the coding scheme developed in Section III to obtain the following subregions of $\mathcal{R}$.

A. A Subregion of $\mathcal{R}$

Let $\mathcal{P}^*$ denote the set of all joint probability distributions $p(\cdot)$ that factors in the form of

$$\begin{aligned} p(q, w, x_1, u, v, \tilde{v}, x_2, y_1, y_2) =& p(q)p(x_1, w|q)p(u|q)p(v, \tilde{v}|w, q) \\ & \cdot p(x_2|u, \tilde{v}, w, q)p(y_1, y_2|x_1, x_2). \end{aligned} \tag{30}$$

Note that the joint distribution (30) differs from (2) in the way that conditioned on Q, U is now independent of any other auxiliary random variables, and $\tilde{U}$ is not present.

Let $\mathcal{R}_{\text{sim}}(p)$ denote the set of all non-negative rate pairs (R_1, R_2) such that

$$R_1 \leq I(W; Y_1|UQ), \tag{31}$$

$$R_2 \leq I(UV; Y_2|Q) - I(V; W|Q), \tag{32}$$

$$R_1 + R_2 \leq I(WU; Y_1|Q) + I(V; Y_2|UQ) - I(V; W|Q); \tag{33}$$

$$0 \leq I(V; Y_1|UQ) - I(V; W|Q), \tag{34}$$

for a joint probability distribution $p(\cdot) \in \mathcal{P}^*$. Furthermore, let

$$\mathcal{R}_{\text{sim}} = \bigcup_{p(\cdot)\in\mathcal{P}^*} \mathcal{R}_{\text{sim}}(p).$$

Theorem 2: The rate region $\mathcal{R}_{\text{sim}}$ is achievable for the discrete memoryless IC-DMS, i.e., $\mathcal{R}_{\text{sim}} \subseteq \mathcal{R} \subseteq \mathcal{C}$.

Proof: The proof can be devised from the proof of Theorem 1 by customizing the original coding scheme for the new joint distribution (30). We change the encoding and decoding method for the message w_{21} (corresponding to U), i.e., the Gel'fand-Pinsker coding used in the proof of Theorem 1 was replaced by the conventional random coding. Specifically, we generate $2^{nR_{21}}$ independent codewords $\mathbf{u}(k_1)$, $k_1 \in \{1, \ldots, 2^{nR_{21}}\}$, according to $\prod_{i=1}^n p(u_i|q_i)$. The encoding and decoding are then adapted to the new codebook accordingly. Evaluating the probability of error in the same way as was done in the proof of Theorem 1, we obtain

$$\tilde{R}_{22} - R_{22} \geq I(V; W|Q); \tag{35}$$

$$R_1 \leq I(W; Y_1|UQ), \tag{36}$$

$$R_1 + R_{21} \leq I(WU; Y_1|Q); \tag{37}$$

$$R_{21} \leq I(U; Y_2|VQ), \tag{38}$$

$$\tilde{R}_{22} \leq I(V; Y_2|UQ), \tag{39}$$

$$R_{21} + \tilde{R}_{22} \leq I(UV; Y_2|Q). \tag{40}$$

Again, we choose $\tilde{R}_{22} - R_{22} = I(V; W|Q)$ in (35), and then substitute $\tilde{R}_{22}$ with $R_{22} + I(V; Y_2|UQ)$ as well as R_{21} with $R_2 - R_{22}$ in the group of (36)–(40). By applying Fourier-Motzkin elimination on the resulting inequalities to remove R_{22}, and adding the constraints that ensure the respective rates R_1, R_{21} and R_{22} to be non-negative, we obtain (31)–(34). Therefore, the region $R_{\text{sim}}(p)$ is achievable for a given $p(\cdot) \in \mathcal{P}^*$, and the theorem follows. ■

Note that simultaneous decoding (simultaneous joint typicality) is applied at both decoders. The advantage of simultaneous decoding over successive decoding is well demonstrated on the IC by Han and Kobayashi in [7]. We next modify the coding scheme by applying successive decoding instead of simultaneous decoding at both decoders to derive a subregion of $\mathcal{R}_{\text{sim}}$.

B. A Subregion of $\mathcal{R}_{sim}$

Let $\mathcal{R}_{\text{suc}}(p)$ denote the set of all achievable rate pairs (R_1, R_2) such that

$$R_1 \leq I(W;Y_1|UQ), \tag{41}$$

$$R_2 \leq \min\{I(U;Y_1|Q), I(U;Y_2|Q)\} + I(V;Y_2|UQ) - I(V;W|Q); \tag{42}$$

$$0 \leq I(V;Y_1|UQ) - I(V;W|Q), \tag{43}$$

for a fixed joint probability distribution $p(\cdot) \in \mathcal{P}^*$. Define

$$\mathcal{R}_{\text{suc}} = \bigcup_{p(\cdot)\in\mathcal{P}^*} \mathcal{R}_{\text{suc}}(p).$$

Theorem 3: The rate region $\mathcal{R}_{\text{suc}}$ is achievable for the discrete memoryless IC-DMS, i.e., $\mathcal{R}_{\text{suc}} \subseteq \mathcal{R}_{\text{sim}} \subseteq \mathcal{R} \subseteq \mathcal{C}$.

Proof: The codebook generation, encoding and transmission remain the same as those used to prove Theorem 2, whereas the decoding processes at both decoders are altered. Both decoders decode w_{21} first, and then decoder 1 decodes w_1 and decoder 2 decodes w_{22} respectively. Then the following can easily be obtained

$$\tilde{R}_{22} - R_{22} \geq I(V;W|Q), \tag{44}$$

$$R_{21} \leq I(U;Y_1|Q), \tag{45}$$

$$R_1 \leq I(W;Y_1|UQ), \tag{46}$$

$$R_{21} \leq I(U;Y_2|Q), \tag{47}$$

$$\tilde{R}_{22} \leq I(V;Y_2|UQ). \tag{48}$$

From (44)–(48), it is straightforward to obtain (41)–(43). Therefore, the region $\mathcal{R}_{\text{suc}}(p)$ is achievable, and the theorem follows immediately. ■

Remark 2: Note that (45) is only necessary when the successive decoding is applied. This is because every decoding step in a successive decoding scheme is expected to have a vanishing probability of error.

In what follows, we further specialize the subregion $\mathcal{R}_{\text{suc}}$ to obtain two more achievable rate regions $\mathcal{R}_{\text{sp1}}$ and $\mathcal{R}_{\text{sp2}}$. Let $\mathcal{P}_1^*$ denote the set of all joint probability density distributions $p(\cdot)$ that factor in the form of

$$p(q,w,x_1,v,\tilde{v},x_2,y_1,y_2) = p(q)p(x_1,w|q)p(v,\tilde{v}|w,q)p(x_2|\tilde{v},w,q)p(y_1,y_2|x_1,x_2). \tag{49}$$

Let $\mathcal{R}_{\text{sp1}}(p)$ denote the set of all non-negative rate pairs (R_1, R_2) such that

$$R_1 \leq I(W;Y_1|Q), \tag{50}$$

$$R_2 \leq I(V;Y_2|Q) - I(V;W|Q), \tag{51}$$

for a fixed joint distribution $p(\cdot) \in \mathcal{P}_1^*$. Define

$$\mathcal{R}_{\text{sp1}} = \bigcup_{p(\cdot)\in\mathcal{P}_1^*} \mathcal{R}_{\text{sp1}}(p).$$

Corollary 1: The region $\mathcal{R}_{\text{sp1}}$ is an achievable rate region for the discrete memoryless IC-DMS, i.e., $\mathcal{R}_{\text{sp1}} \subseteq \mathcal{R}_{\text{suc}} \subseteq \mathcal{R}_{\text{sim}} \subseteq \mathcal{R} \subseteq \mathcal{C}$.

Proof: Fixing the auxiliary random variable U as a constant, we reduce (41) and (42) to (50) and (51), and the corollary follows immediately. ■

Remark 3: We note that the region $\mathcal{R}_{\text{sp1}}$ is similar to the region $\mathcal{R}_{\text{in}}$ reported in [3, Theorem 3.1]. It seems that the region $\mathcal{R}_{\text{in}}$ is more general than the region $\mathcal{R}_{\text{sp1}}$ in the sense that fixing the auxiliary random variable U in $\mathcal{R}_{\text{in}}$ as a constant, one can obtain a region which is the same as $\mathcal{R}_{\text{sp1}}$. Nevertheless, after examining the coding scheme used in [3, Theorem 3.1], one can find that there exists a one-one correspondence between codewords $\mathbf{u}(w_2)$ and $\mathbf{x}_2(w_2)$, and both codewords are jointly generated and decoded, i.e., $p(u, x_2)$ is used to generate two-letter codewords. Thus, one can introduce one auxiliary random variable W such that there exists a one-one mapping between $\mathcal{W}$ and $\mathcal{U} \times \mathcal{X}_2$, i.e., $f : \mathcal{U} \times \mathcal{X}_2 \leftrightarrow \mathcal{W}$, and thus W has the probability mass distribution $p(w) = p(f^{-1}(w)) = p(u, x_2)$. Replacing all $(X_2^n(w_2), U^n(w_2))$ by $W^n(w_2)$ in the proof of [3, Theorem 3.1] will yield the same rate region. Equivalently speaking, for any input distribution achieving a rate region characterized by [3, Theorem 3.1], one can find a corresponding joint distribution in the form of (49) such that Corollary 1 yields exactly the same rate region. Therefore, two rate regions $\mathcal{R}_{\text{sp1}}$ and $\mathcal{R}_{\text{in}}$ are identical.

Let $\mathcal{P}_2^*$ denote the set of all joint probability distributions $p(\cdot)$ that factor in the form of

$$p(q, w, x_1, u, x_2, y_1, y_2) = p(q)p(x_1, w|q)p(u|q)p(x_2|u, w, q)p(y_1, y_2|x_1, x_2). \tag{52}$$

Let $\mathcal{R}_{\text{sp2}}(p)$ denote the set of all non-negative rate pairs (R_1, R_2) such that

$$R_1 \leq I(W; Y_1|UQ), \tag{53}$$

$$R_2 \leq \min\{I(U; Y_1|Q), I(U; Y_2|Q)\}, \tag{54}$$

for a fixed joint distribution $p(\cdot) \in \mathcal{P}_2^*$. Define

$$\mathcal{R}_{\text{sp2}} = \bigcup_{p(\cdot)\in\mathcal{P}_2^*} \mathcal{R}_{\text{sp2}}(p).$$

Corollary 2: The region $\mathcal{R}_{\text{sp2}}$ is an achievable rate region for the discrete memoryless IC-DMS, i.e., $\mathcal{R}_{\text{sp2}} \subseteq \mathcal{R}_{\text{suc}} \subseteq \mathcal{R}_{\text{sim}} \subseteq \mathcal{R} \subseteq \mathcal{C}$.

Proof: The proof can be devised from the proof of Theorem 3 easily by fixing V as a constant. ■

V. The Gaussian IC-DMS

In the preceding sections, we have derived several achievable rate regions for the discrete memoryless IC-DMS. We now extend these results to obtain corresponding achievable rate regions for the *Gaussian* IC-DMS (GIC-DMS).

A. The Channel Model of the GIC-DMS

In general, with no loss of information-theoretic optimality, the GIC-DMS can be converted to the GIC-DMS in the standard form through invertible transformations [2], [6], [11]. We thus only consider the GIC-DMS in the standard form, which is represented as follows

$$Y_1 = X_1 + \sqrt{c_{21}}X_2 + Z_1,$$

$$Y_2 = X_2 + \sqrt{c_{12}}X_1 + Z_2,$$

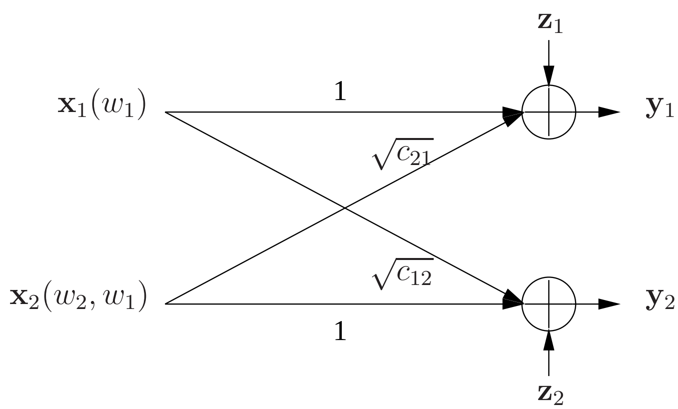

Fig. 4. A Gaussian interference channel with degraded message sets.

where Z_i, $i = 1, 2$, is the additive white Gaussian noise with zero mean and unit variance, and $\sqrt{c_{21}}$ and $\sqrt{c_{12}}$ are the *normalized* link gains in the GIC-DMS depicted in Fig. 4. Moreover, the transmitted codeword $\mathbf{x}_i = (x_{i1}, \ldots, x_{in})$, $i = 1, 2$, is subject to an average power constraint given by

$$\frac{1}{n}\sum_{t=1}^{n} \|x_{it}\|^2 \leq P_i.$$

Since it has been shown in the maximum-entropy theorem in [8] that Gaussian input signals are optimal for Gaussian channels, we will consider Gaussian codewords X_i^n, $i = 1, 2$.

B. Achievable Rate Regions for the GIC-DMS

1) Gaussian Extension of $\mathcal{R}$: We first extend $\mathcal{R}$ to its Gaussian counterpart denoted by $\mathcal{G}$. To obtain the rate region $\mathcal{G}$, we map the random variables involved in the joint distribution (2) to the corresponding Gaussian random variables with the following customary constraints:

P1) W, distributed according to $\mathcal{N}(0, 1)$,
P2) $X_1 = \sqrt{P_1}W$,
P3) $\tilde{U}$, distributed according to $\mathcal{N}(0, \alpha\beta P_2)$,
P4) $\tilde{V}$, distributed according to $\mathcal{N}(0, \alpha\bar{\beta} P_2)$,
P5) $U = \tilde{U} + \lambda_1 W$,
P6) $V = \tilde{V} + \lambda_2 W$,
P7) $X_2 = \tilde{U} + \tilde{V} + \sqrt{\bar{\alpha}P_2}W$,

where $\alpha, \beta \in [0, 1]$, $\alpha + \bar{\alpha} = 1$, $\beta + \bar{\beta} = 1$, $\lambda_1, \lambda_2 \in [0, +\infty)$, and W, $\tilde{U}$ and $\tilde{V}$ are mutually independent. The input-output relationship of the GIC-DMS can be described by

$$Y_1 = \left(\sqrt{P_1} + \sqrt{c_{21}\bar{\alpha}P_2}\right) W + \sqrt{c_{21}}\tilde{U} + \sqrt{c_{21}}\tilde{V} + Z_1, \tag{55}$$

$$Y_2 = \tilde{U} + \tilde{V} + \left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right) W + Z_2. \tag{56}$$

To simplify the derivations, we fix the time-sharing random variable Q as a constant. The issue of how this time-sharing random variable affects the achievable rate region is well addressed

in [13]. In the Gaussian case, the respective mutual information terms in (3) – (9) need be evaluated with respect to the mappings defined by P1–P7. Since the computation procedure to obtain $\mathcal{G}$ and the resulting description of $\mathcal{G}$ are fairly lengthy, we relegate them (Theorem 5) to the Appendix.

2) Gaussian Extension of $\mathcal{R}_{suc}$: For illustration and comparison purpose, we next show how to obtain the Gaussian counterpart of $\mathcal{R}_{\text{suc}}$ in details. Following the first step in the previous derivation, we also map the random variables involved in (30) to the Gaussian ones with the following constraints:

M1) W, distributed according to $\mathcal{N}(0,1)$,
M2) $X_1 = \sqrt{P_1}W$,
M3) U, distributed according to $\mathcal{N}(0,\alpha\beta P_2)$,
M4) $\tilde{V}$, distributed according to $\mathcal{N}(0,\alpha\bar{\beta} P_2)$,
M5) $V = \tilde{V} + \lambda W$,
M6) $X_2 = U + \tilde{V} + \sqrt{\bar{\alpha}P_2}W$,

where $\alpha,\beta \in [0,1]$, $\alpha+\bar{\alpha}=1$, $\beta+\bar{\beta}=1$, $\lambda\in[0,+\infty)$, and W, U and $\tilde{V}$ are mutually independent. Using the mappings defined by M1–M6, we express the input-output relationship for the GIC-DMS as:

$$Y_1 = \left(\sqrt{P_1}+\sqrt{c_{21}\bar{\alpha}P_2}\right)W + \sqrt{c_{21}}U + \sqrt{c_{21}}\tilde{V} + Z_1, \tag{57}$$

$$Y_2 = U + \tilde{V} + \left(\sqrt{\bar{\alpha}P_2}+\sqrt{c_{12}P_1}\right)W + Z_2. \tag{58}$$

Let $\mathcal{G}_{\text{suc}}(\alpha,\beta)$ denote the set of all the non-negative rate pairs (R_1,R_2) such that

$$R_1 \le \frac{1}{2}\log_2\left(1+\frac{\left(\sqrt{P_1}+\sqrt{c_{21}\bar{\alpha}P_2}\right)^2}{c_{21}\alpha\bar{\beta}P_2+1}\right), \tag{59}$$

$$\begin{aligned} R_2 \le& \frac{1}{2}\log_2(1+\alpha\bar{\beta}P_2) + \min\left\{\frac{1}{2}\log_2\left(1+\frac{c_{21}\alpha\beta P_2}{\left(\sqrt{P_1}+\sqrt{c_{21}\bar{\alpha}P_2}\right)^2+c_{21}\alpha\bar{\beta}P_2+1}\right),\right. \\ &\left.\frac{1}{2}\log_2\left(1+\frac{\alpha\beta P_2}{\alpha\bar{\beta}P_2+\left(\sqrt{\bar{\alpha}P_2}+\sqrt{c_{12}P_1}\right)^2+1}\right)\right\}. \end{aligned} \tag{60}$$

Define

$$\mathcal{G}_{\text{suc}} = \bigcup_{\alpha,\beta\in[0,1]} \mathcal{G}_{\text{suc}}(\alpha,\beta).$$

Theorem 4: The region $\mathcal{G}_{\text{suc}}$ is an achievable rate region for the GIC-DMS in the standard form.

Proof: It suffices to prove that $\mathcal{G}_{\text{suc}}(\alpha,\beta)$ is achievable for any given $\alpha,\beta\in[0,1]$. Since $\mathcal{G}_{\text{suc}}$ is extended from $\mathcal{R}_{\text{suc}}$, we need compute the mutual information terms in (41) and (42). The righthand side of (59) can be readily obtained through a straightforward computation of $I(W;Y_1|UQ)$ in (41). Recall that Q is a constant. It is also fairly straightforward to obtain the two terms within the minimum operator in (60) through computing $I(U;Y_1|Q)$ and $I(U;Y_2|Q)$ in (42). We next evaluate the only remaining term $I(V;Y_2|UQ)-I(V;W|Q)$ for a constant Q.

Defining $\tilde{Y}_2 = \tilde{V} + \left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right) W + Z_2$, we have

$$\begin{aligned} I(V;Y_2|U) - I(V;W) &= h(Y_2|U) - h(Y_2|UV) - I(V;W) \\ &= h(\tilde{Y}_2) - h(\tilde{Y}_2|V) - I(V;W) \\ &= h(\tilde{Y}_2) + h(V) - h(\tilde{Y}_2 V) - I(V;W). \end{aligned} \tag{61}$$

With $V = \tilde{V} + \lambda W$, we evaluate (61) as

$$\begin{aligned} & I(V;Y_2|U) - I(V;W) \\ =& \frac{1}{2}\log_2\left(2\pi e\left(\alpha\bar{\beta}P_2 + \left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right)^2 + 1\right)\right) + \frac{1}{2}\log_2(2\pi e(\alpha\bar{\beta}P_2 + \lambda^2)) \\ & - \frac{1}{2}\log_2\left((2\pi e)^2\left[\left(\alpha\bar{\beta}P_2 + \left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right)^2 + 1\right)(\alpha\bar{\beta}P_2 + \lambda^2 P_1)\right.\right. \\ & \left.\left. - \left(\alpha\bar{\beta}P_2 + \lambda\left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right)\right)^2\right]\right) - \frac{1}{2}\log_2\left(1 + \frac{\lambda^2}{\alpha\bar{\beta}P_2}\right). \end{aligned} \tag{62}$$

It is easy to find that when

$$\lambda = \frac{\alpha\bar{\beta}P_2\left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right)}{\alpha\bar{\beta}P_2 + 1}, \tag{63}$$

the term $I(V;Y|U) - I(V;W)$ is maximized, and the maximum value is

$$\max[I(V;Y_2|U) - I(V;W)] = \frac{1}{2}\log_2(1 + \alpha\bar{\beta}P_2). \tag{64}$$

This is in parallel with the result in [12].

Therefore, the rate region $\mathcal{G}_{\text{suc}}(\alpha,\beta)$ is achievable for any pair $\alpha,\beta \in [0,1]$, and the theorem follows. ■

In the following, we obtain two corollaries by setting $\beta = 0$ and $\beta = 1$ in Theorem 4, respectively.

Corollary 3: The rate region $\mathcal{G}_{\text{sp1}}$ is an achievable rate region for the GIC-DMS in the standard form with $\mathcal{G}_{\text{sp1}} := \bigcup_{\alpha\in[0,1]} \mathcal{G}_{\text{suc}}(\alpha,0)$, i.e., $\mathcal{G}_{\text{sp1}}$ is the union of the sets of non-negative rate pairs (R_1,R_2) satisfying

$$\begin{aligned} R_1 &\le \frac{1}{2}\log_2\left(1 + \frac{\left(\sqrt{P_1} + \sqrt{c_{21}\bar{\alpha}P_2}\right)^2}{c_{21}\alpha P_2 + 1}\right), \\ R_2 &\le \frac{1}{2}\log_2(1 + \alpha P_2), \end{aligned}$$

over all $\alpha \in [0,1]$.

Corollary 4: The rate region $\mathcal{G}_{\text{sp2}}$ is an achievable rate region for the GIC-DMS in the standard form with $\mathcal{G}_{\text{sp2}} := \bigcup_{\alpha\in[0,1]} \mathcal{G}_{\text{suc}}(\alpha,1)$, i.e., $\mathcal{G}_{\text{sp2}}$ is the union of the sets of non-negative rate pairs (R_1,R_2) satisfying

$$\begin{aligned} R_1 \le& \frac{1}{2}\log_2\left(1 + \left(\sqrt{P_1} + \sqrt{c_{21}\bar{\alpha}P_2}\right)^2\right), \\ R_2 \le& \min\left\{\frac{1}{2}\log_2\left(1 + \frac{c_{21}\alpha P_2}{\left(\sqrt{P_1} + \sqrt{c_{21}\bar{\alpha}P_2}\right)^2 + 1}\right), \frac{1}{2}\log_2\left(1 + \frac{\alpha P_2}{\left(\sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1}\right)^2 + 1}\right)\right\}, \end{aligned}$$

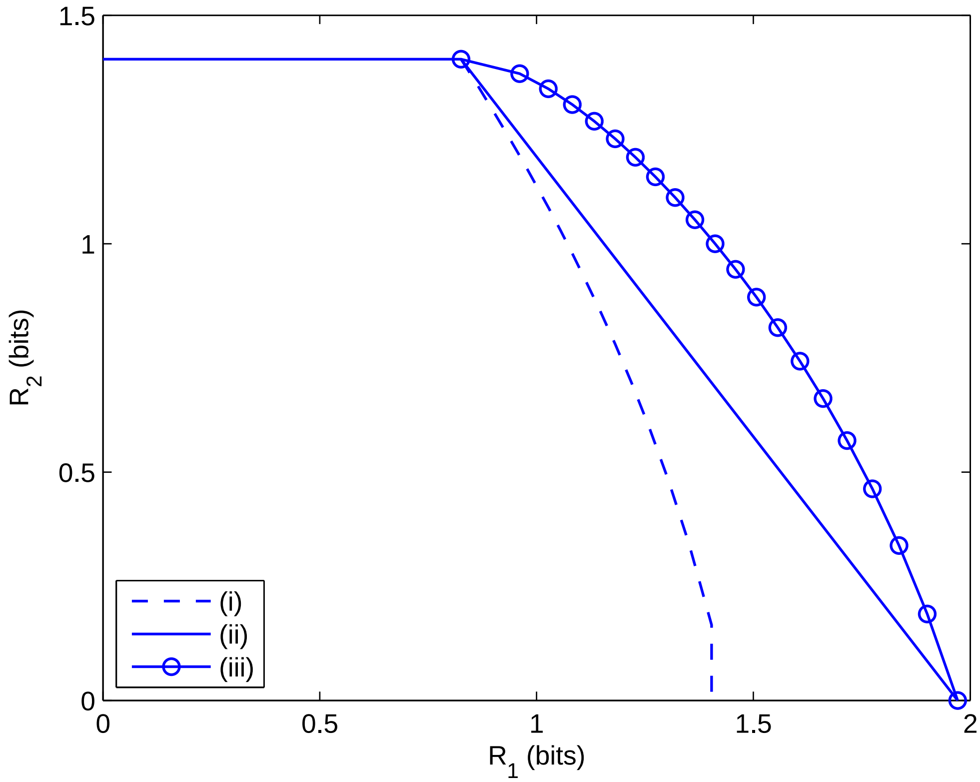

Fig. 5. $P_1 = P_2 = 6$, $c_{21} = 0.3$, $c_{12} = 0$. (i) gives the rate region in Theorem 1 of [1]; (ii) gives the rate region in Corollary 2 of [1]; (iii) gives the rate region in Corollary 3 (equivalently, Theorem 4.1 of [2] and Theorem 3.5 of [3]).

over all $\alpha \in [0, 1]$.

Remark 4: Corollaries 3 and 4 correspond the Gaussian extensions of Corollaries 1 and 2 respectively. Particularly, the rate region depicted by Corollary 3 is the same as the rate regions given in [2, Theorem 4.1] and [3, Theorem 3.5]. It has been proven in both [2] and [3] that the rate region $\mathcal{G}_{\text{sp1}}$ is indeed the capacity region for the GIC-DMS in the low-interference-gain regime, i.e., $c_{21} \leq 1$.

In addition, the set of achievable rate pairs given in [2, Lemma 4.2] is contained in the region $\mathcal{G}_{\text{sp2}}$ as a subset.

C. Numerical Examples

We next provide several numerical examples to illustrate improvements of our achievable rate regions over the previously known results in [1]–[3]. Denote the achievable rate regions obtained in [1, Theorem 1] and [1, Corollary2] by $\mathcal{G}_{\text{dmt1}}$ and $\mathcal{G}_{\text{dmt2}}$, respectively.

1) Comparing with Rate Regions in [1]: Fig. 5 compares the rate regions $\mathcal{G}_{\text{dmt1}}$, $\mathcal{G}_{\text{dmt2}}$, and $\mathcal{G}_{\text{sp1}}$ for an extreme case in which receiver 2 does not experience any interference from sender 1, i.e., $c_{12} = 0$. As can be seen from Fig. 5, the rate region $\mathcal{G}_{\text{sp1}}$ strictly includes $\mathcal{G}_{\text{dmt1}}$, as well as $\mathcal{G}_{\text{dmt2}}$ obtained through time-sharing between $\mathcal{G}_{\text{dmt1}}$ and a fully-cooperative rate point. The

coding scheme used to establish $\mathcal{G}_{\text{dmt1}}$ incurs certain rate loss due to the fact that sender 2 does not use its power to help the sender 1's transmissions even though it has complete and non-causal knowledge about the message being transmitted by sender 1. In contrast, our proposed coding scheme allows sender 2 to use superposition coding to help sender 1, and thus yields an improved rate region.

In Fig. 6, we consider another case in which the transmit power of sender 1 is set to zero and $c_{21} \leq 1$. From the figure, we observe that the rate region $\mathcal{G}_{\text{dmt2}}$ is strictly smaller than $\mathcal{G}_{\text{sp1}}$. Note that in this case, the GIC-DMS becomes a Gaussian degraded broadcast channel. According to [8], the optimal coding scheme for this case is: sender 2 uses a portion of its power to transmit the codeword conveying w_1, and uses the remaining power to transmit the codeword conveying w_2, which is encoded by using the dirty-paper coding [14]. It is easy to verify that this scheme is a special case of the coding scheme developed in Theorem 1.

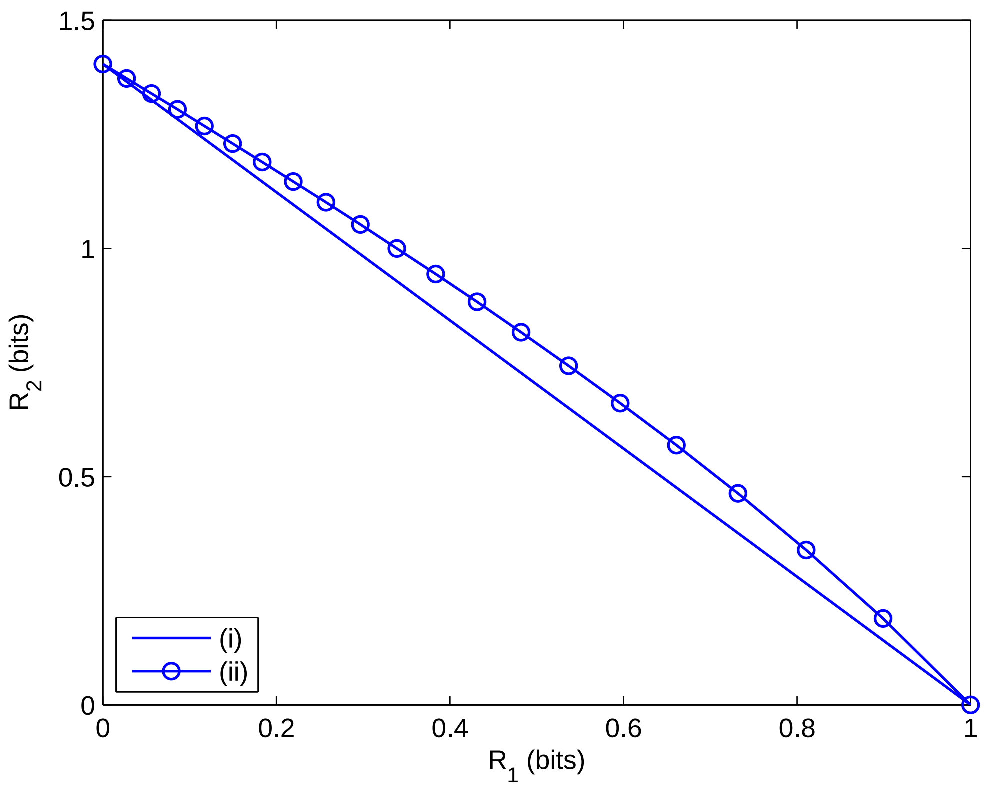

Fig. 6. $P_1 = 0$, $P_2 = 6$, $c_{21} = 0.5$. (i) gives the rate region in Corollary 2 of [1]; (ii) gives the rate region in Corollary 3 (equivalently, Theorem 4.1 of [2] and Theorem 3.5 of [3]).

2) Comparing with Rate Regions in [2], [3]: As mentioned earlier, the rate region $\mathcal{G}_{\text{sp1}}$, a subregion of $\mathcal{G}$, is the same as the one given in [2, Theorem 4.1] and the one given in [3, Theorem 3.5], which is indeed the capacity region for GIC-DMS in the low-interference-gain regime. In Figs. 7 and 8, we compare $\mathcal{G}$ with $\mathcal{G}_{\text{sp1}}$ and $\mathcal{G}_{\text{sp2}}$ in the high-interference-gain regime, i.e., $c_{21} > 1$. As can be seen from the figures, the rate region $\mathcal{G}$ strictly includes both $\mathcal{G}_{\text{sp1}}$ and $\mathcal{G}_{\text{sp2}}$ in this case. Comparing Fig. 7 with Fig. 8, we observe that the improvement of the rate

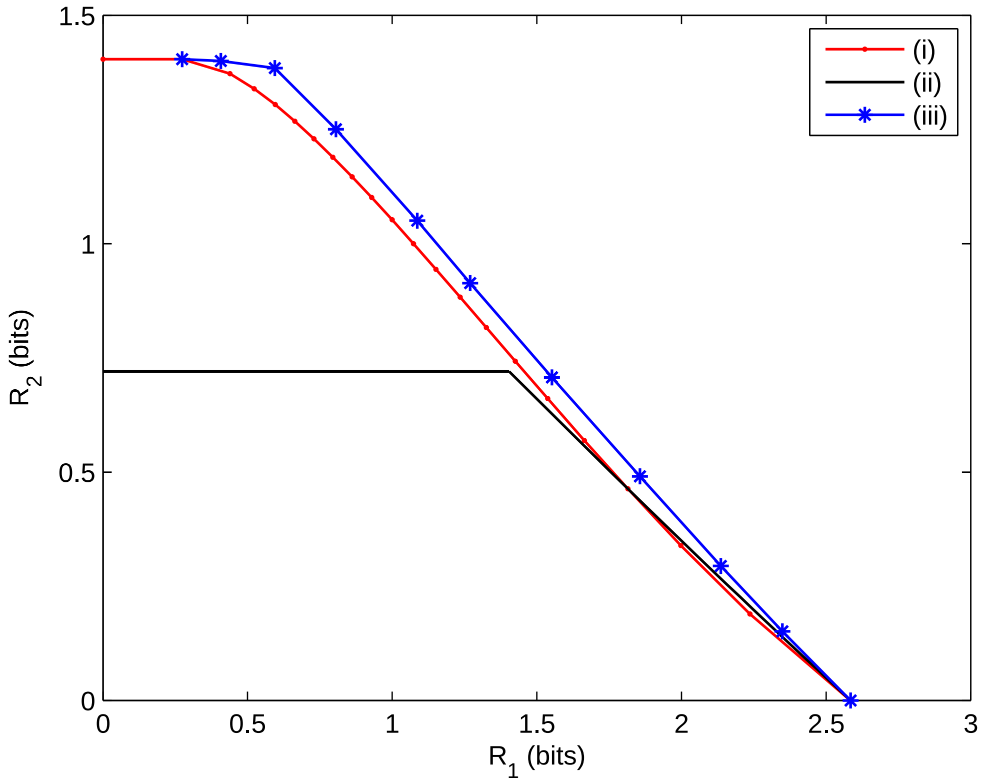

Fig. 7. $P_1 = P_2 = 6$, $c_{21} = 2$, $c_{12} = 0.3$. (i) gives the rate region in Corollary 3 (equivalently, Theorem 4.1 of [2] and Theorem 3.5 of [3]); (ii) gives the achievable rate region in Corollary 4; (iii) gives the achievable rate region in Theorem 5.

region $\mathcal{G}$ over $\mathcal{G}_{\text{sp1}}$ becomes more pronounced as the link gain c_{21} increases. The improvement is mainly because our coding scheme allows receiver 1 to decode partial information from sender 2, and thus reduces the effective interference experienced by receiver 1. In addition, it can be seen from the figures that in the high-interference-gain regime, $\mathcal{G}_{\text{sp1}}$ is not convex and thus is only suboptimal.

VI. Conclusions

In this paper, we have investigated the IC-DMS from an information theoretic perspective. We have developed a coding scheme that combines the advantages of cooperative coding, collaborative coding and Gel'fand-Pinsker coding. With the coding scheme, we have derived a new achievable rate region for such a channel, which not only includes existing results as special cases, but also exceeds them in the high-interference-gain regime. However, we are not able to establish a converse for the derived achievable rate region, because the achievable result is closely related to the achievable results for the interference channel and the broadcast channel, for which there is no converse available in general.

Appendix

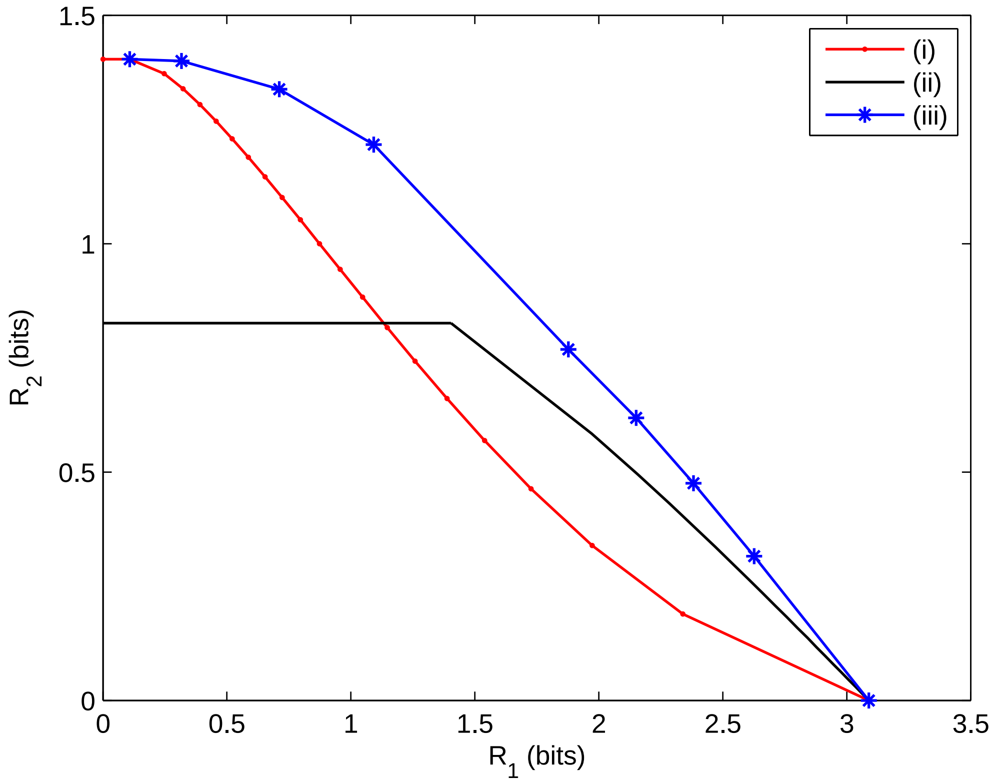

Fig. 8. $P_1 = P_2 = 6$, $c_{21} = 6$, $c_{12} = 0.3$. (i) gives the rate region in Corollary 3 (equivalently, Theorem 4.1 of [2] and Theorem 3.5 of [3]); (ii) gives the achievable rate region in Corollary 4; (iii) gives the achievable rate region in Theorem 5.

An Achievable Rate Region for the GIC-DMS

In this appendix, we show how to extend $\mathcal{R}$, the achievable rate region for the discrete memoryless IC-DMS, to its Gaussian counterpart, $\mathcal{G}$. Note that the mappings M1–M6 of the auxiliary random variables are described in Section V. We first compute the following two covariance matrices:

$$\begin{aligned}\mathbf{\Sigma}_{WUY_1} &= \begin{pmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{pmatrix} := \begin{pmatrix} E\{W^2\} & E\{WU\} & E\{WY_1\} \\ E\{WU\} & E\{U^2\} & E\{UY_1\} \\ E\{WY_1\} & E\{UY_1\} & E\{Y_1^2\} \end{pmatrix} \\ &= \begin{pmatrix} P_1 & \lambda_1 P_1 & \eta_1\sqrt{P_1} \\ \lambda_1 P_1 & \alpha\beta P_2 + \lambda_1^2 P_1 & \lambda_1\eta_1\sqrt{P_1} + \sqrt{c_{21}}\alpha\beta P_2 \\ \eta_1\sqrt{P_1} & \lambda_1\eta_1\sqrt{P_1} + \sqrt{c_{21}}\alpha\beta P_2 & \eta_1^2 + c_{21}\alpha P_2 + 1 \end{pmatrix},\end{aligned}$$

$$
\begin{aligned}
\mathbf{\Sigma}_{UVY_2} &= \begin{pmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{pmatrix} := \begin{pmatrix} E\{U^2\} & E\{UV\} & E\{UY_2\} \\ E\{UV\} & E\{V^2\} & E\{VY_2\} \\ E\{UY_2\} & E\{VY_2\} & E\{Y_2^2\} \end{pmatrix} \\
&= \begin{pmatrix} \alpha\beta P_2 + \lambda_1^2 P_1 & \lambda_1\lambda_2 P_1 & \alpha\beta P_2 + \lambda_1\eta_2\sqrt{P_1} \\ \lambda_1\lambda_2 P_1 & \alpha\bar{\beta}P_2 + \lambda_2^2 P_1 & \alpha\bar{\beta}P_2 + \lambda_2\eta_2\sqrt{P_1} \\ \alpha\beta P_2 + \lambda_1\eta_2\sqrt{P_1} & \alpha\bar{\beta}P_2 + \lambda_2\eta_2\sqrt{P_1} & \alpha P_2 + \eta_2^2 + 1 \end{pmatrix},
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_1 &= \sqrt{P_1} + \sqrt{c_{21}\bar{\alpha}P_2}, \\
\eta_2 &= \sqrt{\bar{\alpha}P_2} + \sqrt{c_{12}P_1},
\end{aligned}
$$

and $E\{\cdot\}$ denotes the expectation of a random variable.

Define $\Gamma(x) = \log_2(x)/2$, and $\xi = \log_2(2\pi e)/2$. We express the respective differential entropy terms as:

$$
\begin{aligned}
h_a &= h(W) = \xi + \Gamma(\mu_{11}), \\
h_b &= h(UY_1) = 2\xi + \Gamma\left(\begin{vmatrix} \mu_{22} & \mu_{23} \\ \mu_{32} & \mu_{33} \end{vmatrix} \right), \\
h_c &= h(WUY_1) = 3\xi + \Gamma\left(\begin{vmatrix} \mu_{11} & \mu_{12} & \mu_{13} \\ \mu_{21} & \mu_{22} & \mu_{23} \\ \mu_{31} & \mu_{32} & \mu_{33} \end{vmatrix} \right), \\
h_d &= h(UV) = 2\xi + \Gamma\left(\begin{vmatrix} \nu_{11} & \nu_{12} \\ \nu_{21} & \nu_{22} \end{vmatrix} \right), \\
h_e &= h(Y_2) = \xi + \Gamma(\nu_{33}), \\
h_f &= h(UVY_2) = 3\xi + \Gamma\left(\begin{vmatrix} \nu_{11} & \nu_{12} & \nu_{13} \\ \nu_{21} & \nu_{22} & \nu_{23} \\ \nu_{31} & \nu_{32} & \nu_{33} \end{vmatrix} \right), \\
h_g &= h(WU) = 2\xi + \Gamma\left(\begin{vmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{vmatrix} \right), \\
h_h &= h(Y_1) = \xi + \Gamma(\mu_{33}), \\
h_i &= h(V) = \xi + \Gamma(\nu_{22}), \\
h_j &= h(UY_2) = 2\xi + \Gamma\left(\begin{vmatrix} \mu_{11} & \mu_{13} \\ \mu_{31} & \mu_{33} \end{vmatrix} \right), \\
h_k &= h(U) = \xi + \Gamma(\nu_{11}), \\
h_l &= h(VY_2) = 2\xi + \Gamma\left(\begin{vmatrix} \mu_{22} & \mu_{23} \\ \mu_{32} & \mu_{33} \end{vmatrix} \right),
\end{aligned}
$$

where $|\cdot|$ denotes the determinant of a matrix.

The mutual information terms in (3)–(9) are then computed as:

$$I_1 = h_a + h_b - h_c,$$
$$I_2 = h_d + h_e - h_f,$$
$$I_3 = \Gamma(1 + \frac{\lambda_1^2 P_1}{\alpha\beta P_2}),$$
$$I_4 = \Gamma(1 + \frac{\lambda_2^2 P_1}{\alpha\bar{\beta} P_2}),$$
$$I_5 = h_g + h_h - h_c,$$
$$I_6 = h_i + h_j - h_f,$$
$$I_7 = h_k + h_l - h_f.$$

Let $\mathcal{G}(\alpha, \beta, \lambda_1, \lambda_2)$ denote the set of all rate pairs (R_1, R_2) such that the following inequalities are satisfied:

$$R_1 \leq I_1, \tag{65}$$
$$R_2 \leq I_2 - I_3 - I_4, \tag{66}$$
$$R_1 + R_2 \leq I_5 + I_6 - I_3 - I_4; \tag{67}$$
$$0 \leq I_5 - I_3, \tag{68}$$
$$0 \leq I_7 - I_3, \tag{69}$$
$$0 \leq I_6 - I_4, \tag{70}$$
$$0 \leq I_2 - I_3 - I_4. \tag{71}$$

for given $\alpha, \beta \in [0, 1]$ and $\lambda_1, \lambda_2 \in [0, +\infty)$. Note that (65)–(71) are directly extended from (3)–(9).

Theorem 5: The rate region $\mathcal{G}$ is achievable for the GIC-DMS in the standard form with

$$\mathcal{G} = \bigcup_{\alpha,\beta\in[0,1];\lambda_1,\lambda_2\in[0,+\infty)} \mathcal{G}(\alpha, \beta, \lambda_1, \lambda_2).$$

References

[1] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[2] A. Jovičić and P. Viswanath, "Cognitive radio: An information-theoretic perspective," 2006. [Online]. Available: http://arxiv.org/abs/cs/0604107

[3] W. Wu, S. Vishwanath, and A. Arapostathis, "On the capacity of interference channels with degraded message sets," 2006. [Online]. Available: http://arxiv.org/abs/cs/0605071

[4] I. Maric, R. Yates, and G. Kramer, "The strong interference channel with unidirectional cooperation," in *Proc. UCSD Workshop on Information Theory and its Applications*, San Diego, CA, USA, Feb. 2006.

[5] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[6] A. B. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[7] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1994.

[9] T. M. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 399–404, July 1975.

[10] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications, ser. CISM Courses and Lectures,G. Longo, Ed. Springer-Verlag*, vol. 229, 1978, pp. 171–231.

[11] G. Kramer, "Review of rate regions for interference channels," in *Proc. IZS Workshop*, Zurich, Feb. 2006.

[12] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[13] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 1345–1356, June 2004.

[14] W. Yu and J. Cioffi, "Trellis precoding for the broadcast channel," in *Proc. IEEE Global Telecommunications Conference (Globecom)*, San Antonio, TX, Nov. 2001.